DOCUMENT RESUME

ED 135 855 TM 006 086

AUTHOR Blue, Terry W.
TITLE The Effect of Written and Oral Student Evaluative Feedback and Selected Teacher and Student Demographic and Descriptive Variables on the Attitudes and Ratings of Teachers and Students.
PUB DATE [Apr 77]
NOTE 32p.; Paper presented at the Annual Meeting of the American Educational Research Association (61st, New York, New York, April 4-8, 1977)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS Changing Attitudes; *Feedback; Group Relations; Junior High Schools; School Attitudes; Self Evaluation; Statistical Analysis; *Student Attitudes; Student Characteristics; *Student Evaluation of Teacher Performance; *Student Teacher Relationship; *Teacher Attitudes; Teacher Characteristics

ABSTRACT

Nine hypotheses dealing with the effect of four modes of student evaluative feedback on eight measures of teacher and student attitudes and ratings were tested. Five ancillary questions were investigated to assess the effect of teacher and student demographic variables, group processes variables, the application of attitude change theory, and the interrelationship of measures. A consistent pattern was found on all comparisons of the effects of treatment on eight posttest measures. A number of teacher and student demographic variables, selected group processes variables, and level of teacher-student rating discrepancy were found to significantly influence attitude and rating scores. Suggestions for the improvement of teaching behaviors, student behaviors, and learning environments were offered. Implications for teacher education were presented. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED135855

The Effect of Written and Oral Student Evaluative Feedback and Selected Teacher and Student Demographic and Descriptive Variables on the Attitudes and Ratings of Teachers and Students

by
Terry W. Blue, Ph.D.
Chairman and Assistant Professor
Department of Education
Franklin and Marshall College
Lancaster, Pennsylvania 17604

A paper presented at the 1977 Annual Meeting
of the American Educational Research Association,
New York, New York, April 5, 1977, Session 5:11

TM006 086

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The Effect of Written and Oral Student Evaluative Feedback and Selected Demographic and Descriptive Variables on the Attitudes and Ratings of Teachers and Students

TERRY W. BLUE

Franklin and Marshall College

Nine hypotheses dealing with the effect of four modes of student evaluative feedback on eight measures of teacher and student attitudes and ratings were tested. Five ancillary questions were investigated to assess the effect of teacher and student demographic variables, group processes variables, the application of attitude change theory, and the interrelationship of measures. A consistent pattern was found on all comparisons of the effects of treatment on eight posttest measures. A number of teacher and student demographic variables, selected group processes variables, and level of teacher-student rating discrepancy were found to significantly influence attitude and rating scores. Suggestions for the improvement of teaching behaviors, student behaviors, and learning environments were offered. Implications for teacher education were presented.

INTRODUCTION

It is expected that teachers should be constantly working to improve their teaching skills and to develop stronger, more effective means of communicating cognitive and affective skills to their students. The improvement of teaching behaviors, however, usually requires some procedure to help the teacher to identify those actions which are effective in encouraging viable learning experiences for students (Bebb, Low, and Waterman, 1969). This is the role that modern supervision has attempted to fill.

Supervision in the public schools, however, is going through a continuous process of change. Over the years many systems of behavioral observation have been developed and used in the classroom (Simon and Boyer, 1967). Each system assumes that behavior, observed over a given period of time, is a meaningful measure from which interpretations can be made. Most of these systems utilize an outside observer(s) whose investment in the particular class setting is a minimal one. Supervision additionally is often irregular and incomplete. Finally, if teaching is to improve, the teacher must have continuous access to reliable information about his effectiveness. Clear-cut proof of success is difficult for a teacher to get, and the uncertainty has profound effects on attitudes and conduct (Snow, 1963).

In his review of the research on the effects of teacher behaviors on student achievement, Rosenshine (1971) has reported that general student ratings have a particularly good history in offering such input to teachers. Others have concurred in that view (Remmers, 1960; Bannister, 1961; Tuckman and Oliver, 1968; Centra, 1974). Recent years have seen an increase of interest in the topic supporting the feeling that various types of student evaluation and feedback seem to

be useful, convenient, reliable, and valid means of self-supervision and self-improvement for teachers.

Many demographic and descriptive variables, however, have been found to confound studies of student evaluation of teachers (Yee, 1968, 1970; Paulk, 1972; Williams, 1973). Others (Solomon, Rosenberg, and Bezdek, 1964; Rayder, 1968); Kerlinger, 1963) have shown that students' perceptions of teachers are more a function of internal frames of reference or value systems than of concrete teacher characteristics. The students attitudes and values are therefore of much concern. The same can be said of teachers' attitudes and values as well. Smith (1971) has offered that there can be little doubt that the attitudes teachers have toward themselves and their classes influence their behavior in the classroom.

The multiplexity of both teacher and student attitudes further complicates the picture of the classroom as a social system (Krech, Crutchfield, and Ballachey, 1962). Bigelow (1971) proposed that such attitudes color all aspects of classroom behavior. Schmuck (1971) further supported Bigelow's assertion when he suggested that informal features of classroom groups often have important bearing on formal aspects. Such features require continued investigation.

If the quality of teaching and learning is to improve, many changes remain to be made. Miles and Schmuck (1971) have suggested that schools tend to hire and retain people with dependent, submissive attitudes who have a difficult time in situations requiring the exercise of open, frank problem solving. This, they offered, explains the low rate of instructional change characteristic of education. What is perhaps needed is what McGregor (1961) termed the new vision of man in the organization. Instead of Theory X, which viewed man as indolent, passive,

self-protective, and requiring managerial control from above, schools should be searching for, or better yet be creating, Theory Y qualities which picture man as inherently curious and capable of growth, of being trustworthy, and of taking initiative. Any attempt to empirically establish viable means for teachers to move toward Theory Y qualities seems to require support.

STATEMENT OF THE PROBLEM

The major problem examined in this study was whether daily and weekly written student evaluative feedback to teachers, weekly teacher-student evaluative discussions, or daily and weekly written student evaluative feedback with weekly teacher-student evaluative discussions, and the resultant teacher self-evaluation stimulated by each method would cause a change in:

1. students' attitudes toward self
2. students' attitudes toward others
3. students' attitudes toward the school environment
4. students' ratings of their teachers
5. teachers' attitudes toward self
6. teachers' attitudes toward others
7. teachers' attitudes toward the school environment
8. teachers' self-ratings

Another principal concern of the study was to determine the effect of the various pretest measures as viable treatments by themselves: that is, to measure the effect of pretest sensitization on the areas above.

Secondary concerns of the investigation included attempts to determine if differences in posttest scores on teacher and student attitude and rating measures could be explained by:

1. teacher and student demographic variables, such as teacher experience, level of professional preparation, sex, and subject assignment and student sex, grade range, and involvement in extra-curricular activities

2. group processes variables, such as cohesiveness, peer liking structure, leadership, norm, and communication patterns

3. the application of attitude change theory, particularly the idea of discrepancy between expected and realized ratings for teachers

4. inter-relationships of the four measures used for both teachers and students---attitudes toward self, attitudes toward others, attitudes toward the school environment, and student and teacher ratings and self-ratings

HYPOTHESES AND ANCILLARY QUESTIONS

The following null hypotheses were tested:

1. There will be no significant differences in posttest scores caused by treatment on attitudes toward self for students as measured by the About Myself Scale on the STS Youth Inventory.

2. There will be no significant differences in posttest scores caused by treatment on attitudes toward self for teachers as measured by the Acceptance of Self Scale of the Berger Instrument.

3. There will be no significant differences in posttest scores caused by treatment on attitudes toward others for students as measured by the Getting Along with Others Scale of the STS Youth Inventory.

4. There will be no significant differences in posttest scores caused by treatment on attitudes toward others for teachers as measured by the Acceptance of Others Scale of the Berger Instrument.

5. There will be no significant differences in posttest scores caused by treatment on attitudes toward the school environment for students as measured by the My School Scale of the STS Youth Inventory.

6. There will be no significant differences in posttest scores caused by treatment on attitudes toward the school environment for teachers as measured by the Purdue Teacher Opinionaire.

7. There will be no significant differences on posttest scores caused by treatment on ratings of teachers by students as measured by the Purdue Teacher Evaluation Scale.

8. There will be no significant differences in posttest scores caused by treatment on self-ratings of teachers as measured by the Purdue Teacher Evaluation Scale.

9. There will be no significant differences in posttest scores caused by treatment between the student ratings of teachers and the self-ratings of these same teachers as measured by the Purdue Teacher Evaluation Scale.

The following questions were examined to provide peripheral data related to the major hypotheses of this study:

1. Are there any significant differences caused by teacher experience level, subject assignment, level of professional preparation, sex, age, grade assignment, or involvement in extracurricular activities on the reported posttest scores of teacher and student attitude measures, teacher and student rating measures, and student descriptive variables?

2. Are there any significant differences caused by student sex, grades received, involvement in extracurricular activities, ability grouping, or plans after high school on the reported posttest scores of teacher and student attitude measures, teacher and student rating measures, and student demographic and descriptive variables?

3. Are there any significant differences caused by class cohesiveness, peer liking structure, norms, leadership, or communication patterns on the reported posttest scores of teacher and student attitude measures, teacher and student rating measures, and student demographic and descriptive variables?

4. Are there any significant differences caused by the level of discrepancy of student ratings and teacher self-ratings on the Purdue Teacher Evaluation Scale on the reported posttest scores of teacher attitude and rating measures, student attitude and rating measures, and student descriptive variables?

5. Are there any significant differences caused by teacher or student level on one posttest measure on the reported posttest scores of the other three attitude or rating measures?

EXPERIMENTAL PROCEDURES

Design of the study

The design used for this study was an adaptation and extension of the Solomon Four-Group design (Campbell and Stanley, 1963). This design is graphically presented by the following:

R'	O_1	X_1	O_2	(Group 1)
R'		X_1	O_3	(Group 2)
R'	O_4	X_2	O_5	(Group 3)
R'		X_2	O_6	(Group 4)
R'	O_7	X_3	O_8	(Group 5)
R'		X_3	O_9	(Group 6)
R'	O_{10}		O_{10}	(Group 7)
R'			O_{12}	(Group 8)

In this design the symbol R' denotes the stratified random assignment of teachers and intact class units of students into the eight groups of the study. The symbol of O represents the pretest and posttest measures while the symbol X denotes the treatment utilized. The subscripts used with the symbol O were merely to identify these elements for statistical analysis. The subscripts used with X further describe the treatment with X_1 representing the treatment consisting of both daily and weekly written student evaluation and feedback and weekly teacher-student evaluative discussion, X_2 representing the treatment consisting of daily and weekly written student evaluation and feedback only, and X_3 representing the treatment consisting of weekly teacher-student evaluative discussion only. Groups 7 and 8 in this design acted as control groups for the treatment in the study, while Groups, 2, 4, 6, and 8 acted as control groups for pretesting.

Selection of subjects

The teacher and student sample used in this study was drawn from a junior high school in Pennsylvania. Total student population of this school was approximately 1300 students in grades seven, eight, and nine. Average class size in this school ranged from twenty to thirty with all classes used in the study being homogeneously grouped.

The total number of teachers in this school during the experimental period was seventy-two with fifty-two qualifying for inclusion in the study by meeting the conditions established in the design. Of these fifty-two teachers, forty-eight were randomly selected so that all experimental groups would be of equal size. At the completion of the study, one teacher per experimental group had to be excluded from consideration as several teachers submitted incomplete data. Forty teachers (five in each group) and their classes were therefore included in the study.

Teachers in the study were randomly assigned to each of the eight groups by grade. That is, all groups had an equal number of teachers from any one grade, two, yielding a total teacher population in each group of six. This balance was subsequently lost when teachers presenting incomplete data were excluded. All teachers in the school who met with the same students daily were assigned to one of the groups so far as it was mathematically possible. Teachers were therefore used from the English, reading, mathematics, science, social studies, and foreign language departments.

Student assignment into groups was on the basis of intact classes determined by their teacher's group placement. Confounding of student groups through interaction with teachers of differing groups was avoided by choosing one period per grade level when all students were assigned to their major teachers and using this period as the class used with each teacher. The median class size used in the study was twenty-four students.

Background data for use in analysis of ancillary questions were collected for teachers in the sample including: sex, age, level of professional preparation, teaching experience level, grade and subject placement, and involvement in extracurricular activities. Descriptive information for students was collected including: number going on to various types of additional study after high school, sex, grades received, ability grouping, subject they were rating, and involvement in extracurricular activities.

Pretest and posttest instrumentation

Pretest and posttest instruments for students in the study included:

1. elements of the STS Youth Inventory dealing with My School, About Myself, and Getting Along with Others

2. the Purdue Teacher Evaluation Scale, a student rating of teachers' ability to motivate, ability to control, subject matter orientation, communication patterns, teaching methods and procedures, and fairness

Pretest and posttest instruments for teachers in the study included:

1. the Purdue Teacher Opinionaire, a teacher morale measure made up of ten factors including: rapport with principal, teaching satisfaction, rapport among teachers, salary, teaching load, curriculum issues, status, community support, school facilities, and community pressure

2. an adapted version of the Purdue Teacher Evaluation Scale used as a self-rating.

3. The Berger Instrument of Acceptance of Self and Others

Group processes instruments

Cohesiveness of each class was measured using an adaptation of a teacher rating scale of group cohesiveness developed by Bany and Johnson (1964).

Peer liking structure of individual classes was determined using a procedure described by Schmuck (1963).

Norm patterns of each class were measured using an investigator-designed instrument based on the work of Jackson (1960).

Leadership patterns of each class were determined using a procedure similar to that utilized for peer liking structure (Schmuck, 1963).

Communication patterns for each class were measured using a technique described by Schmuck (1971).

Feedback instruments and discussion model

Instruments used for daily and weekly evaluation and feedback to treatment teachers were created by the investigator. Each week's instrument was designed to

be increasingly sophisticated and comprehensive when compared to the previous week's measure. These instruments were largely based on examples offered by Schmuck (1968), Medley and Mitzel (1963), Remmers (1963), Ryans (1960), Rosenshine (1971), Rosenshine and Furst (1971), Schmuck, Runkel and Langmeyer (1971), and Schmuck and Schmuck (1971).

The model used for weekly teacher led discussions was created by the investigator based on the pattern described above for the feedback instruments.

Experimental treatment

Treatment for Groups 1 and 2 included:

1. daily and weekly written student evaluation and feedback using a checklist format developed by the investigator. Comments were entered on a daily basis on this form and were submitted weekly to the teacher involved

2. brief weekly discussions (approximately ten minutes) of these evaluations between the teacher and his/her class following guidelines provided by the investigator

Treatment for Groups 3 and 4 included:

1. daily and weekly written student evaluation and feedback using a checklist format developed by the investigator. Comments entered on this form on a daily basis and were submitted weekly to the teacher involved. No discussion of the feedback followed, however

Treatment for Groups 5 and 6 included:

1. brief weekly discussions (approximately ten minutes) between teachers and students of the activities of their class in the past week following discussion guidelines provided by the investigator

Groups 7 and 8 did not receive treatment of this type, but instead maintained their normal classroom activities during the experimental period.

Experimental methodology

The experimental aspect of this study began at the start of the fourth quarter of the school year. This gave students and teachers an opportunity to become well acquainted and allowed the attitudes of both to reach the stable level described by Tuckman and Oliver (1968), Wong (1971), and Flanders, Morrison, and Brode (1970). During this waiting period the experiment was organized, materials were prepared, and groups were structured.

Pretest data were collected for students in Groups 1, 3, 5, and 7 during a two day period. Tests were administered to these students in the academic sections making up these four groups. Teachers in Groups 1, 3, 5, and 7 also were given the three pretest measures to complete during this time. All pretest data were submitted to the experiment supervisor upon completion. A Student Questionnaire and Teacher Rating Scale were completed during the first week of the experiment and were also submitted to the supervisor upon completion.

When all pretest data were collected, experimental treatment began. At the start of each of the six weeks of the study, teachers in Groups 1, 2, 3, and 4 distributed a checklist type of evaluative device to their students. This instrument was divided into six sections. Five of the sections were used for daily evaluation while the remaining section was used to record weekly evaluative responses. The content of the instrument changed every two weeks, gradually moving toward more sophisticated commentary and evaluation by students. Teachers gave students a few minutes each day to complete these forms. At the end of each week the completed form was turned over to the teacher.

Teachers in Groups 1 and 2 then spent approximately ten minutes of a subsequent class meeting discussing these anonymous evaluations with their students, following a format provided by the investigator. Teachers in Groups 3 and 4

were asked to refrain from discussing the evaluations they collected, putting off, if necessary, student attempts to do so. This pattern was followed for each of the six weeks of the study for these groups.

Teachers in Groups 5 and 6 meanwhile were carrying out only the second phase described above. They used the discussion outline provided by the investigator to lead their classes in a weekly discussion of the past week's lessons with no prior evaluation by students.

Teachers in Groups 7 and 8 continued to conduct their classes in their normal manner during the six weeks of the study.

At the end of the six week experimental period, all teachers were directed to suspend all evaluative exchanges. Posttests were administered to all teachers and students in all eight groups following the procedures described above for pretesting. After all the posttest data were collected, teachers in all groups were allowed to resume or begin student evaluation if they so desired. An anecdotal experiment questionnaire was completed by teachers at this time and the results of the experiment were discussed.

Control of treatment

In order to insure that treatment took place as designed, teachers in Groups 1, 2, 3, and 4 were required to send their completed student evaluation forms to an experiment supervisor to be recorded on a weekly basis. Once the completion of this phase of the study was certified by the supervisor, the forms were returned to the teachers for their future use. Each week teachers in Groups 1, 2, 5, and 6 were required to submit a brief summary of the class discussions they led to the supervisor. These, too, were returned to teachers once their completion was certified and recorded.

Teachers were additionally asked to maintain and certify a log indicating the dates each of the activities required of them and their classes relative to this study was completed. This log was collected at the close of the experimental period.

Finally, teachers were asked to avoid discussing the activities of their experimental class relative to this study with teachers of other groups.

One additional related consideration merits reporting. A series of one-factor analyses of variance were run on the pretest means of teachers and students in an effort to determine if any significant differences existed in the initial scores of such participants. No significant differences among pretested groups were found on any teacher or student score including the six subscales of the PTES. Consideration of the main effect of pretesting on posttest scores will be treated below.

ANALYSIS OF DATA

Nine hypotheses in null form and five ancillary questions were postulated relative to this study.

Two-factor analysis of variance was used to test each of the nine hypotheses. The main effects of pretesting, treatment, and the interaction of pretesting and treatment were determined. One-factor analysis of variance and the Tukey WSD technique were used to further identify any significant differences that were found. A significance level of .05 was used in each of these tests. Trends were, however, identified at higher levels.

A series of one-factor analyses of variance was used to answer four of the five ancillary questions. The fifth ancillary question was evaluated through a series of Pearson Product-Moment Correlations. A .10 significance level was employed on these tests of the ancillary questions because of their secondary, descriptive role in the study.

CONCLUSIONS

Based on a literal interpretation and analysis of data presented in Chapter IV, the following conclusions relative to this study can be offered:

1. Pretest sensitization was not a significant cause of differences on seven of the eight measures used in this study. Only in the measurement of student attitude toward school was there a significant main effect of pretesting

2. Table 1 presents a summary of significant differences, trends, and order means from analyses of variance and pair-wise contrasts on the hypotheses. In each of the twenty posttest measures (including the subscales of the PTES) the means of teachers and students who participated in Treatment 3 (discussion only) were higher than or more positive than those of Treatment 2 (written feedback only), Treatment 1 (both written and oral feedback), and Treatment 4 (the control group). In three instances (student attitude toward self and teacher self-ratings on Scales 5 and 6 of the PTES) the differences were significant at the .05 level. These significant differences were between Treatments 3 and 1 on student attitude toward self and Scale 5 of the PTES and between Treatments 3 and 4 on Scale 6 of the PTES. In tests where differences were not significant at the designated level, a pattern nonetheless was evidenced. This pattern saw Treatment 3 surpassing Treatment 1 on virtually all posttest measures. Means for Treatments 4 and 2 generally fell between those for Treatments 3 and 1, but in varying order. It would appear then from a literal interpretation of the test results that teacher-student evaluative discussion was the most effective means of student evaluation of teachers tested in this study. The strength of this pattern was most intense on student attitude and teacher self-rating measures, less intense on teacher attitude measures, and least intense on student rating measures.

Table 1. Summary of Significant Differences, Trends, and Order of Means from Analyses of Variance and Pair-Wise Contrasts on Hypotheses

Hypoth.	Measure	Significant Differences	Trends[a]	Order of Means
1	STSYI-Myself[b]	1 > 3 F=2.943 @ 2.72[c]	1 > 4 F=1.638 @ 2.72	1 > 2 > 4 > 3
3	STSYI-Others[b]	n.s.d.[d]	1 > 3 @ .083[e]	1 > 4 > 2 > 3
5	STSYI-School[b]	n.s.d.	1 > 3 @ .142	1 > 2 > 4 > 3
7	PTES-Student			
	Total	n.s.d.	3 > 1 @ .419	3 > 4 > 2 > 1
	Scale 1	n.s.d.	3 > 1 @ .356	3 > 4 > 2 > 1
	Scale 2	n.s.d.	3 > 2 @ .748	3 > 4 > 1 > 2
	Scale 3	n.s.d.	3 > 1 @ .649	3 > 4 > 2 > 1
	Scale 4	n.s.d.	3 > 1 @ .332	3 > 4 > 2 > 1
	Scale 5	n.s.d.	3 > 1 @ .328	3 > 4 > 2 > 1
	Scale 6	n.s.d.	3 > 2 @ .351	3 > 1 > 4 > 2
2	Berger-Self	n.s.d.	3 > 1 @ .090	3 > 2 > 4 > 1
4	Berger-Others	n.s.d.	3 > 1 @ .318	3 > 2 > 4 > 1
6	PTO	n.s.d.	3 > 2 @ .165	3 > 1 > 4 > 2
8	PTES-Teacher			
	Total	n.s.d.	3 > 1 @ .078	3 > 2 > 4 > 1
	Scale 1	n.s.d.	3 > 1 @ .227	3 > 2 > 4 > 1
	Scale 2	n.s.d.	3 > 1 @ .446	3 > 2 > 4 > 1
	Scale 3	n.s.d.	3 > 1 @ .492	3 > 1 > 4 > 2
	Scale 4	n.s.d.	3 > 1 @ .358	3 > 2 > 4 > 1
	Scale 5	3 > 1 F=3.055 @ 2.72	2 > 1 F=2.626 @ 2.72 4 > 1 F=1.876 @ 2.72	3 > 2 > 4 > 1
	Scale 6	3 > 4 F=3.272 @ 2.72	2 > 4 F=2.004 @ 2.72 1 > 4 F-1.803 @ 2.72	3 > 2 > 1 > 4
9	PTES-Teacher/ Student	n.s.d.	3 > 1 @ .060	3 > 2 > 4 > 1

a differences not significant at .05
b lower score = more positive value
c critical t value
d no significant differences
e probability

3. When teacher posttest means were grouped by various levels on teacher demographic variables, a number of significant influences were isolated. Teacher self-ratings on the PTES were found to be significantly related to teacher sex, professional preparation, experience, age, subject and grade taught. Teacher attitude scores were found to be significantly influenced by teacher levels on sex, age, activities, and subject variables. Student attitudes and teacher ratings were less subject to such influences. Only teacher levels on sex and preparation were significantly responsible for differences in student ratings while only teacher levels on experience and grade variables significantly influenced student attitude scores. Finally, teachers' sex, preparation, experience, activities, and subject levels were found to be significantly related to student scores on group processes variables

4. The effect of various student demographic variables on teacher posttest scores was less revealing. No significant differences on either student ratings or teacher attitude measures were found to be caused by student level on any of these variables. Some strong probabilities just beyond significance were isolated, however. Student attitude scores were found to be significantly influenced by students' grades and activities while teacher self-ratings were found to be significantly related to student levels on grades, activities, plans, and ability. Strong significant relationships were additionally found among the various demographic variables and the reported scores of each level on the group processes variables

5. When class units were grouped into levels on the basis of scores on group processes variables, strong patterns of relationship were again found. Teacher attitude scores were significantly influenced by students' levels on the cohesiveness, norm, and leadership variables. Self-ratings of teachers were found to be significantly related to students' levels on the cohesiveness, leader-

ship, and peer liking variables. While levels on only the leadership variable offered significant predictions of student ratings of teachers, levels on both the cohesiveness and norm variables yielded influences on student attitude scores. Strong interrelationships of the various group processes variables were also found. In addition, cohesiveness, peer liking, and leadership patterns were found to be significantly related to student demographic characteristics

6. The analysis of the discrepancy between teacher and student posttest scores on the PTES proved an efficient predictor of significant relationship on both teacher and student rating measures. It was not, however, found to be as effective on either teacher or student attitude measures

7. When each of the eight posttest measures was correlated with each of the seven other measures, significant relationships were found on all but four of the possible combinations (student-others and teacher-school, student-school and teacher-school, student PTES and teacher-PTES, and teacher-school and teacher-self). After the direction of the STS Youth Inventory scores was reversed, six of the negative or inverse relationships were associated with student attitude toward others. An improvement (decrease) on such scores led to a decrease in student attitude toward self and school, in teacher attitude toward self and others, and in teacher and student rating means. The final negative relationship was found between student ratings and teacher attitude toward the school environment

DISCUSSION OF THE CONCLUSIONS

While the statistical analysis and interpretation of data which resulted in the above conclusions represented the legitimate and best effort of the investigator in assessing the hypotheses of this study as they were designed, an element of concern nonetheless remained after their presentation. This element

of concern dealt primarily with the conclusions of this study relative to the hypotheses.

While no significant differences were found among pretested groups of any of the eight posttest measures and while only one significant effect of pretesting was found among posttest means in the consideration of the eight attitude and rating measures, post hoc analysis of the ancillary questions of this investigation indicates that in the future control over the four treatment groups could be improved.

Through the structure provided by the first three ancillary questions, several demographic and group processes variables were found to have a significant influence on teacher and student attitude and rating scores. Tables 2 and 3 present a summary of those variables by treatment and an indication of the level of influence each had on attitude and rating scores. The tables additionally show the direction of influence by treatment. In every case where a variable had high influence on attitude and rating measures, the direction of influence favored Treatment 3 over the other treatments. In eight of the ten high influence variables Treatments 1 and 4 held the two lowest positions on the direction of influence. It should be indicated that while no significant differences were found for any of the variables when treatment means were compared, the advantage gained by Treatment 3 when all high influence variables are combined would certainly seem to be an influence in explaining the differences found in the analysis of the first eight hypotheses. It is therefore suggested that conclusion two above be amended by the following:

2. .....The above effects of treatment were possibly confounded by the high influence that certain teacher demographic, student demographic, and group processes variables were found to have on teacher and student attitude and rating measures. The combination of these influences seemed to have favored

Table 2. Summary of Teacher Demographic Variables Showing Level and Direction of Influence and Treatment Benefitted by Each Variable

Variable	Category	Treatment 1	2	3	4	Level of Influence	Direction of Influence	Treatment Benefitted
Sex	Male	5	4	3	5	High	F>M	3>2>1&4
	Female	5	6	7	5			
Age	20-25	3	5	2	4	Moderate	H>M&L	3>1>4>2
	26-35	6	3	4	5			
	Over 35	1	2	4	1			
Preparation	B to B+29	7	7	3	5	High	H>L	3>4>1&2
	Over B+29	3	3	7	5			
Experience	0-2 Yrs.	1	2	2	3	Low	H>M>L	3>1>2>4
	3-7 Yrs.	6	5	3	5			
	Over 7 Yrs.	3	3	5	2			
Grade	Seven	4	4	4	2	Moderate	Irregular	Irregular
	Eight	3	4	3	4			
	Nine	3	2	3	4			
Subject	Soc. St.	1	2	1	2	Moderate	Irregular	Irregular
	Lang. Arts	2	3	5	3			
	Math	5	2	0	1			
	Science	1	2	3	3			
	For. Lang.	1	1	1	1			
Extra-curricular Activities	Zero	2	3	3	3	Moderate	L&N>H&M	3>4>2>1
	One	2	2	6	2			
	Two	3	2	1	5			
	Over Two	3	3	0	0			

Table 3. Summary of Student Demographic and Group Processes Variables Showing Level and Direction of Influence and Treatment Benefitted by Each Variable

Variable	Treatment 1	2	3	4	Level of Influence	Direction of Influence	Treatment Benefitted
Grades	4.838	4.931	5.223	4.599	High	H>M>L	3>2>1>4
Plans	3.204	3.320	3.525	3.172	High	H>M>L	3>2>1>4
Activities	3.800	3.218	3.587	2.975	Moderate	Irregular	Irregular
Sex Ratio	.939	1.348	.910	2.235	Moderate	L>M>H	3>1>2>4
Ability	3.6	2.9	3.1	3.2	Moderate	Irregular	Irregular
N	23.1	23.4	25.0	22.2	High	H>M>L	3>1&2>4
Cohesion	43.2	46.9	53.5	43.2	High	H>M>L	3>2>1&4
Communication	3.123	3.307	3.274	3.323	Low	H>L>M	2&4>1>3
Norms	23.66	23.66	23.99	23.70	Low	H&M>L	3&4>1&2
Peer Liking Liked	.621	.561	.464	.567	High	D>M>B	3>2&4>1
Peer Liking Disliked	.465	.455	.377	.464	High	D>M>B	3>2>1&4
Leadership Leaders	.494	.448	.366	.485	High	D>M>B	3>2>1&4
Leadership Non-Leaders	.483	.421	.404	.523	High	D>M>B	3>2>1>4

Treatment 3 over the other treatments in the facilitation of positive attitude and rating change.

The other conclusions would remain as presented above.

IMPLICATIONS

It was the goal of this study to extend the research on this topic and to combine a number of variables in a novel and comprehensive manner. The conclusions offered above would seem to indicate some measure of success in defining the effects of various forms of evaluative feedback from students to teachers, in clarifying the effects of various teacher and student demographic variables and student group processes variables, in applying an attitude change theory to the process of student and teacher attitude change, in measuring the effect of pretest sensitization on teacher and student attitude and rating measures, and in determining the relationship among various attitude and rating measures. As a result of these findings, some suggestions for those interested in improving the quality of educational experiences seem in order.

Since the relationship between certain group processes variables and teacher and student attitude and rating measures was found to be so strong, it would seem that all educators should be more cognizant of the role such variables play in fixing the social and emotional climate of the classroom. This would seem particularly important in the junior high school setting. If class cohesiveness, peer liking structure, and leadership patterns of classes are such a strong influence on the attitudes and ratings of teachers and students, attempts should be made to encourage high levels of cohesiveness and diffuse patterns of peer liking and leadership in the school and classroom environment. Much more than this, these high cohesion levels and diffuse peer liking and leadership

patterns should be positively related to the goals of the school. Students could, for example, be highly cohesive in their hatred of the teacher and/or school. This type of counterproductive cohesiveness could be discouraged through an openness in relationships between teachers and students, between students and other students, and between teachers and other teachers. The logical extension of this idea then would be to include the administrative team in a series of similar relationships. In the same manner, an openness in these relationships and a high level of regard for the variety of unique capabilities held by all individuals involved in the schools should encourage the diffuse peer liking and leadership patterns so conducive to positive change. The movement toward the Theory Y image of man (McGregor, 1961) in the school environment should be stimulated by such a shift.

A second suggestion would be related to the findings of this study relative to teacher and student demographic variables. If the interests of all classes in a school are to be considered, it would seem imperative that the individual(s) responsible for scheduling in that school attempt to balance the experiences of students with teachers of the various demographic levels. Without such a balance, the opportunities for some students would greatly outweigh those of other students. At the same time, it would seem that the effects of the influence of certain student demographic variables would make a strong case in favor of random assignment of students to classes for the same reasons. Equality of educational opportunity would seem to be limited by more traditional grouping patterns. Development of the high level of cohesiveness and the diffuse patterns of peer liking and leadership should, however, be a necessary, though difficult, concommitant part of such a shift. Without such a change, the mere shift in grouping patterns would probably be of limited value.

Some conflicting suggestions would have to be made based upon the data collected in the study relative to the relationships among the various posttest attitude and rating measures. While in most cases improvement on one measure predicted improvement on the other measures, an opposite pattern was isolated when student attitudes toward others were considered. With the exception of teacher attitudes toward the school environment, where the relationship was not significant, an improvement in student attitudes toward others led to a decrease on the other measures. On the basis of these results it would seem counter-productive to encourage strong student acceptance of others rather than strong acceptance of self. The reason for this pattern remains unclear and requires future examination. The reason for a lack of a relationship between teacher and student scores on the Purdue Teacher Evaluation Scale also remains unclear at this point.

A fourth set of implications relative to this study could be based on the data relative to treatment differences. With the qualifications proposed above, Treatment 3 teachers and students were found to have scored significantly higher on several posttest measures. The encouragement of the discussion format for student feedback to teachers would seem to be indicated. Centra (1974), however, has suggested that teachers who initially rated themselves more favorably than students rated them tended to change in the direction suggested by students. As all of the pretested student groups rated their teachers more unfavorably than did teachers themselves, the application of Centra's theory to the posttest scores on the PTES would indicate that Treatments 1 and 2, which shifted most in their pretest to posttest scores, were the most responsive teacher groups. Treatments 3 and 4 showed little change from pretest to posttest (for those groups who were pretested) and could be said to have changed the least because of

student feedback. This condition would certainly be easily accepted for Treatment 4, the control group.

The change in attitude scores for both teachers and students can be less easily explained however. Little change was found on any measure, but what change there was favored Treatment 3 rather than Treatment 1. The explanation offered earlier, that the confusion and disequilibrium caused by the rigors of the treatment was responsible for this increase in student problems, remains the only suggestion on this concern. A replication with structures built in to measure the direction as well as the magnitude of change would seem in order.

The reactions of teachers involved in the study through their responses on a questionnaire provided by the investigator included some other considerations for future efforts in this research area. Generally teachers and students entered the study with limited enthusiasm for the rather rigorous procedures required. Few had any experience with the whole process of student evaluation. Those teachers who did, had collected data from students only at the end of the year. Many teachers felt that the particular class chosen for them to work with was not an ideal choice, that the time of the year chosen for the study was less than optimal, and that it would have been better if all of their classes had been involved. While many felt that student evaluation was of only limited value after their experience, the majority did not plan to give up thoughts of trying it at another time.

The teachers had strong feelings regarding how the results of student evaluation should be used. Few wanted to share student opinions with even their department chairman much less with a building or central administrator. The strong majority additionally felt that student evaluations should not be employed in determining teacher competence, compensation, and related uses.

From this brief summary one can see that student evaluation of teachers was far from a popular procedure. It was only grudgingly considered for future use. Until teachers can be more open in their dealings with their students, fellow teachers, and curricular leaders, it would seem that chances for significant change toward the Theory Y teacher (McGregor, 1961) are yet small.

RECOMMENDATIONS

Based on the results of this study, it is recommended that this study be expanded with the following changes and/or under the following conditions:

1. The size of teacher and class groups should be enlarged to further increase the power of statistical analysis

2. The population sample utilized should be broadened to include teachers and particularly students of a wider range of socioeconomic status

3. The age/grade level tested should be extended to include both high school and elementary school students and their teachers.

4. The experimental period should be extended to include one school year so that procedures need not be so intense and demanding and so that trends in student and teacher attitudes and ratings over the course of a school year can be plotted

5. Various other forms and frequencies of evaluations should be tested

6. Either no one involved in the study should be pretested (as it was found to have little or no effect) or everyone involved should be pretested (so that the direction of change could be ascertained for all participants)

7. Treatment groups should be controlled for the significant group processes and demographic variables isolated in this study or a design calling for multiple regression analysis should be added

8. If at all possible, permission for access to data on individual students should be secured so that changes in individual student's attitudes and ratings could be determined

9. Measurement of the direction as well as the magnitude of teacher and student change should be attempted

10. More accurate instruments for the measurement of student norm and communication patterns should be developed and utilized

11. Responses of teachers should be additionally analyzed on the basis of their rated effectiveness

12. Other measurements of student attitude toward others should be included to help clarify the relationship of this variable and the other posttest measures

It is further recommended that the scope of this study be extended to include the following:

1. Measurement of the effect of a series of supportive inservice programs for teachers coordinated with the written and oral evaluative feedback from students should be attempted

2. The measurement of the effect of various types of intervention and guidance by administrators, consultants, and other teachers on teachers' attitudes toward the handling of student evaluative feedback should be attempted

3. The determination of the behavioral change caused by evaluative feedback to teachers as measured on both cognitive and affective dimensions by systematic observation instruments should be attempted

4. Finally, it is recommended that any design or procedure that would lead to the encouragement and development of more open relationships within and among classrooms and schools be supported as a prime research focus for American schools.

REFERENCES

Bannister, J. et.al. Evaluating college teaching. Curriculum Reporter Supplement, 1961, 1, 21-25.

Bany, M. A. and Johnson, L. V. Classroom group behavior. New York: the Macmillan Company, 1964.

Bebb, A. M., Low, A. F., and Waterman, F. T. Supervisory conference as individualized teaching. Washington, D. C.: The Association for Student Teaching, 1969.

Bigelow, R. C. Changing classroom interaction through organization development. In Schmuck, R. A. and Miles, M. B. (Eds.), Organization development in schools. Palo Alto, California: National Press Books, 1971, 71-85.

Campbell, D. T. and Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally and Company, 1963.

Centra, J. A. The utility of student ratings for instructional improvement. Princeton: Educational Testing Service, 1974.

Flanders, N. A., Morrison, B. M. and Brode, E. L. Changes in pupil attitudes during the school year. In Purnell, R. F. (Ed.), Adolescents and the American high school. New York: Holt, Rinehart and Winston, 1970, 295-303.

Jackson, J. M. Structural characteristics of norms. In Henry, N. B. (Ed.), The dynamics of instructional groups, 59th Yearbook, Part 2. Chicago: National Society for the Study of Education, 1960, 136-163.

Kerlinger, F. N. Educational attitudes and perceptions of teachers: Suggestions for teacher-effectiveness research. The School Review, 1963, 71, 1-11.

Krech, D., Crutchfield, R. S., and Ballachey, E. L. Individual in society. New York: McGraw-Hill, 1962.

McGregor, D. M. The human side of enterprise. New York: McGraw-Hill, 1961.

Medley, D. M. and Mitzel, H. E. Measuring classroom behavior by systematic observation. In Gage, N. L. (Ed.), Handbook of research on teaching. Chicago: Rand McNally and Company, 1963, 247-328.

Miles, M. B. and Schmuck, R. A. Improving schools through organization development: An overview. In Schmuck, R. A. and Miles, M. B. (Eds.), Organization development in schools. Palo Alto, California: National Press Books, 1971, 1-27.

Paulk, L. J. A study of teacher self-actualization philosophical orientation and student attitudes. Unpublished doctoral dissertation, University of Southern Mississippi, 1972.

Rayder, N. F. College student ratings of instructor. Journal of Experimental Education, 1968, 37, 76-81.

Remmers, H. H. Manual, the Purdue Rating Scale for Instruction. West Lafayette, Indiana: University Book Store, 1960.

Remmers, H. H. Rating methods in research on teaching. In Gage, N. L. (Ed.), Handbook of research on teaching. Chicago: Rand McNally and Company, 1963, 329-378.

Rosenshine, B. Teaching behaviors and student achievement. Windsor, England: National Foundation for Educational Research in England and Wales, 1971.

Rosenshine, B. and Furst, N. Research on teacher performance criteria. In Smith, B. O. (Ed.), Research in teacher education: A symposium. Englewood Cliffs, New Jersey: Prentice-Hall, 1971, 39-58.

Ryans, D. G. Characteristics of teachers. Washington, D. C.: American Council of Education, 1960.

Schmuck, R. A. Some relationships of peer liking patterns in the classroom to pupil attitudes and achievement. The School Review, 1963, 71, 337-359.

Schmuck, R. A. Helping teachers improve classroom group processes. Journal of Applied Behavioral Science, 1968, 4, 401-435.

Schmuck, R. A. Improving classroom group processes. In Schmuck R. A. and Miles, M. B. (Eds.), Organization development in schools. Palo Alto, California: National Press Books, 1971, 29-50.

Schmuck, R. A., Runkel, P., and Langmeyer, D. Using group problem-solving procedures. In Schmuck, R. A. and Miles, M. B. (Eds.), Organization development in schools. Palo Alto, California: National Press Books, 1971, 51-69.

Schmuck, R. A. and Schmuck, P. A. Group processes in the classroom. Dubuque, Iowa: William C. Brown Company, 1971.

Simon, A. and Boyer, E. G. Mirrors for behavior, an anthology of observation instruments. Philadelphia: Research for Better Schools, 1967.

Smith, B. O. (Eds.). Research in teacher education: A symposium. Englewood Cliffs, New Jersey: Prentice-Hall, 1971.

Snow, R. H. Anxieties and discontents in teaching. Phi Delta Kappan, 1963, 44, 318-321.

Solomon, D., Rosenberg, L., and Bezdek, W. E. Teacher behavior and student learning. Journal of Educational Psychology, 1964, 55, 23-30.

Tuckman, B. W. and Oliver, W. F. Effectiveness of feedback to teachers as a function of source. Journal of Educational Psychology, 1968, 59, 297-301.

Williams, D. L. An assessment of factors related to student opinion of school climate. Unpublished doctoral dissertation, Northern Illinois University, 1973.

Wong, I. The effect of pupil feedback concerning student teacher verbal behavior on selected teacher characteristics as perceived by pupils. Unpublished doctoral dissertation, Michigan State University, 1971.

Yee, A. H. Source and direction of casual influence in teacher-pupil relationships. Journal of Educational Psychology, 1968, 59, 275-282. (a)

Yee, A. H. Interpersonal attitudes of teachers and advantaged and disadvantaged pupils. The Journal of Human Resources, 1968, 3, 327-345. (b)

Yee, A. H. Do principals' interpersonal attitudes agree with those of teachers and pupils? Educational Administration Quarterly, 1970, 6, 1-13.